LBNL-60749

Dose-Response Modeling for Life Cycle Impact Assessment: Findings of the Portland Review Workshop

Thomas E. McKone[1,2], Amy D. Kyle[1], Olivier Jolliet[3],
Stig Irving Olsen[4], Michael Hauschild[4]

[1]University of California
School of Public Health
Berkeley, California, 94720, USA

[2]Lawrence Berkeley National Laboratory
Environmental Energy Technologies Division
Indoor Environment Department
Berkeley, California, 94720, USA

[3]University of Michigan
Ann Arbor, Michigan, 48109 USA

[4]Technical University Denmark (DTU)
DK-2800, Lyngby, Denmark

June 2006

This work was supported in part by the US Environmental Protection Agency National Exposure Research Laboratory through Interagency Agreement #DW-89-93058201-1, through the US Department of Energy under Contract Grant No. DE-AC02-05CH11231.

Dose-Response Modeling for Life Cycle Impact Assessment: Findings of the Portland Review Workshop

Thomas E. McKone[1,2], Amy D. Kyle[1], Olivier Jolliet[3], Stig Irving Olsen[4], Michael Hauschild[4]

[1]University of California, School of Public Health, Berkeley, California, 94720, USA
[2]Lawrence Berkeley National Laboratory, Berkeley, California, 94720, USA
[3]University of Michigan, Ann Arbor, Michigan, 48109, USA
[4]Technical University Denmark (DTU), DK-2800, Lyngby, Denmark

With contributions from the workshop participants[1].

Corresponding author: Thomas E. McKone, temckone@lbl.gov

Introduction

The United Nations Environment Program (UNEP)/SETAC Life Cycle Initiative aims at putting life cycle thinking into practice and at improving the supporting tools for this process through better data and indicators. The initiative has thus launched three programs with associated working groups (see http://www.uneptie.org/pc/sustain/lcinitiative/). The *Task Force on Toxic Impacts* was established under the Life Cycle Impact Assessment (LCIA) program to establish recommended practice and guidance for use in human toxicity, ecosystem toxicity, and related categories with direct effects on human health and ecosystem health.

During the 2004 SETAC Europe meeting in Prague, an international group of LCIA practitioners initiated activities of the LCIA Task Force 3 (TF3) to address exposure and toxicity with the goal of establishing guidance for LCIA. As an adjunct activity of the 2004 SETAC World Congress in Portland, Oregon, TF3 members organized a workshop to review existing proposals on human toxicity indicators for LCIA. The particular focus of this workshop was on options regarding dose-effect response and severity. The review workshop consisted of formal presentations of approaches followed by a review discussion performed by a panel of internationally recognized dose-response modeling experts. This workshop was organized by Thomas McKone of the University of California, Berkeley and Michael Hauschild and Stig Irving Olsen from the Danish Technical University in Denmark. Amy Kyle of the University of California, Berkeley, facilitated the workshop. This workshop involved several internationally recognized dose-response modeling experts as well as LCIA specialists (see footnote for the full list). The product of this workshop is a set of short recommendations that are being transmitted via this report.

[1] Besides the authors, other workshop participants were: Lois Swirsky Gold, University of California, Berkeley, CA USA, Lorenz Rhomberg, Reviewer, Gradient Corporation, Cambridge, MA USA, Glenn Suter, US Environmental Protection Agency, Cincinnati, OH, USA, Jane Bare, US Environmental Protection Agency, Thomas Gloria, Five Winds International, Stefanie Hellweg, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, Allan Astrup Jensen, Force Technology, Denmark, Randy Maddalena, Lawrence Berkeley National Laboratory, Guido Sonnemann, Division of Technology, Industry & Economics, United Nations Environment Programme, Paris, France, Dik van de Meent, Member, RIVM Laboratory for Ecological Risk Assessment, Bilthoven, The Netherlands.

Workshop Format

The workshop consisted of three elements.

(A) presentations summarizing (1) the goals of the LCIA Task Force (2) historical approaches to exposure and toxic impacts in LCIA (3) current alternative proposals for addressing human health impacts. Viewgraphs from two of these presentations are provided in Appendix B to this report.

(B) Discussion among a panel of experts about the scientific defensibility of these historical and proposed approaches in the context of the goals of the LCIA Task Force 3 on toxicity impacts.

(C) Development of the recommendations to the LCIA program and working group for optimum short- and long-term strategies for addressing human health impacts in LCA.

Background and Key References

Life cycle assessment (LCA) is a framework for comparing products (or product-related emissions) according to their total estimated environmental impact summed over all chemical emissions and activities associated with the product's life cycle. To assess human toxicity impact, the LCIA practitioner considers for each chemical involved the cumulative exposure associated with the mass released to a defined (indoor, urban, regional, etc.) environment by multiplying the release amount by a measure of toxic impact to characterize the likelihood of health effects and their potential consequences.

The SETAC Life Cycle Impact Assessment (LCIA) Working Group on Human Toxicity (Krewitt et al., 2002) have classified measures of toxic impact into two broad categories: (1) potency-based characterization factors that are used to assess the likelihood of a disease or effect (cancer, death, reproductive failure, etc) and (2) severity-based characterization factors or damage factors that, in addition to the qualitative or quantitative likelihood of disease, reflect population consequences of the disease in terms of years of life loss or some other measure of societal impact. But rather than indicating the likelihood of disease, potency mainly indicates a dose that has an effect (the effect dose or ED) or the lower confidence on the ED (Gold *et al.*, 2003). Recently, Crettaz *et al.* (2002) and Pennington *et al.* (2002) have proposed for both cancer and non-cancer health impacts in LCIA alternative approaches that include elements of both the potency and severity impact measures note above. For carcinogens, Crettaz *et al.* (2002) based their approach on the maximum likelihood estimate of the dose inducing a 10% response over background (the ED10) and derive from this a linear low-dose extrapolation using the slope 0.1/ED10. They obtain the ED10 values by using the US EPA Integrated Risk Information System (IRIS), the medium tumor dose rate (TD50) from the Cancer Potency database of Gold *et al.* (2006), or data on median single lethal dose (LD50). Potential consequences and severity are addressed by combining the low-dose slope with a measure of disability adjusted life years lost due to cancer to obtain a aggregate number of disability adjusted life years (DALYs) attributable to a specified chemical release. For non-carcinogens Pennington *et al.* (2002) follow an approach similar to that of Crettaz *et al.* (2002), but made use of the benchmark dose to obtain a low-dose slope factor for non-cancer diseases.

Review Questions

The workshop organizers provided a set of review questions in advance to both the experts and other workshop participants. With regard to dose-response modeling in LCIA, the workshop organizers structured these review questions into five categories (Table 1)—general issues, measures of potency, species and population extrapolation, measures of severity, and data quality and availability.

Table 1: Questions addressed by the review workshop

1. General issues
What is the scientific evidence for classifying any substance as a human carcinogen?
What is the scientific evidence for classifying any substance as inducing some type of non-cancer disease impact such as neurotoxicity, reproductive toxicity, respiratory irritation, developmental delay, asthma, autoimmune diseases, etc.?
What is the validity and utility of using non-threshold linear dose-response models for assessing either cancer or non-cancer population disease burden?
Are disability adjusted life years (DALYs) an appropriate endpoint measure of disease burden for life-cycle impact?
Are there alternatives to DALYs that should be considered?
Should we consider individual or population (collective) risk as a measure of impact in life-cycle assessment?
When combining the effects from exposures to multiple harmful substances, should we consider effect additivity, risk additivity, or cumulative disease burden in a measure of human health impact for LCA?
2. Measures of potency
What is the basis for the potency measure (e.g. critical effect, most severe effect, etc.)?
Should carcinogens and non-carcinogens be separated when calculating health impact midpoints and/or health impact endpoints?
For non-cancer outcomes should the midpoint and/or endpoint measure of disease burden make use of no adverse/lowest adverse effect (NOAEL/LOAEL) doses or benchmark doses (e.g. TD50, ED10)?
For cancer outcomes should the midpoint and/or endpoint be based on slope factors or other potency measures such as benchmark dose?
3. Species and population extrapolation
In developing dose-response and disease burden models, should we apply extrapolation factors to account for differences between animals and humans, between sensitive subpopulations, etc.?
If we use extrapolation factors, how should the extrapolation factors be derived, e.g. probabilistic methods or some set of default assumptions?

4. Measures of severity

Should and how can LCIA models of toxicity include the severity of the effect?

Should severity assessment be qualitative (indicator-based), semi-quantitative, or quantitative?

If the measure of severity is quantitative: should the model include full damage modeling, that is both morbidity and mortality, and how should mortality and morbidity be aggregated?

Should the model include partial damage modeling (e.g. only mortality)?

What methods of damage modeling should be used?

If the measure of severity is semi-quantitative or qualitative, what approach should be used to categorize and/or classify severity?

5. Data quality and availability

Can and how should data quality for human health impact and disease burden be quality assured?

Are both experimental data and derived "safe" levels such as allowable daily intake (ADI) and reference dose (RfD) from the EPA IRIS database etc. available for a broad set of chemicals?

What is the minimum level of data availability required to establish the lethal dose for 50% of a population (LD50), NOAEL, and/or benchmark doses from chronic studies?

Workshop Findings

The workshop findings are organized into three categories—overarching issues, measures of potency, and measures of severity.

Overarching Issues:

In concurrence with other workshop participants, the dose-response experts concluded that it is appropriate to include human toxicity in the LCIA process. The basis for this recommendation is that, in the absence of a toxicity metric, many LCIA practitioners will continue to rely on emissions magnitude as a measure of emissions impact. But, because of the significant differences among chemicals in the dose levels that are toxic, it is essential to consider human toxicity in comparing releases of different toxic chemicals.

There is an overarching concern that the LCIA process should address toxicity but limit the level of detail in the analysis to information that provides benefits or value to the overall LCIA. Too much detail can reduce the transparency and reliability of the LCIA. However, there may be cases where more analysis matters. That is, cases where exposure is below a threshold of effect and for which more details on the distribution of population exposure will impact estimates of disease burden.

Measures of Potency

The expert panel recommended the use of a hierarchy of toxicity values in LCIA with priorities assigned so that LCIA assessors can evaluate the relative advantages of different toxicity metrics. The goal is to allow LCIA assessors to determine which metrics would be scientifically defensible and informative as well as identifying methods that are transparent and easy to use. With regard to measures of potency, the workshop participants made the following specific recommendations.

a) *Human preference:* Relevant and applicable human data should be given priority wherever such data are available.

b) *Preference for benchmark measures:* The review panel preferred the use of benchmark measures of effect as a means of scaling relative toxicity, rather than a no observed adverse effect level (NOAEL), lowest observed adverse effect level (LOAEL) or Reference Dose (RfD), which is derived from NOAEL or LOAEL. Benchmark measures include the ED10, the dose that results in a toxic effect to 10% of the exposed population, and the ED50, the dose that results in a toxic effect to 50% of the exposed population. The basis for this recommendation is the frequent dependence of the NOAEL or LOAEL on the limitations of study design. However, the experts expressed the view that it would be better to use the NOAEL or LOAEL rather than no measure at all, as a basis to estimate ED10 or ED50. Given the overall uncertainties, there are large differences between these alternate approaches in ranking chemicals in LCIA. RfD values are obtained from NOAELs or LOAELs using safety or uncertainty factors that reflect differing degrees of precaution or protection. As a result, it is not clear whether the RfD, which was developed to provide an adequate margin of safety, can provide the consistent measure of harm needed in LCIA. Thus, when an RfD is used in chemical ranking, the panel recommended that the uncertainty measures used and the corresponding NOAEL or LOEL should be separated out from the RfD and reported along with the RfD.

As a follow-up issue, the panel identified the need to come to a conclusion about whether an ED10 value, an ED50 value, some other benchmark (ED25), or some approach that combines these benchmark doses would be best suited to LCIA. The ED50 is at the median of the range of doses that produce significant results. The advantages of the ED50 are that it is a more stable measure and represents the point most comparable among biological species. The ED10 is usually within the dose range tested for statistically significant results. The advantage of the ED10 is that the information about slope or dose-response that is most chemical specific is between the ED50 and ED10 range, and this is better captured in the ED10. But the use of ED10 will not automatically reduce the level of uncertainty for low dose extrapolation. If a practitioner believes that “slope” values from animal studies predict human values, this would argue for using the ED10. But this is a question that requires further research and for which we do not yet know the answer.

A related need is to determine how to make use of combinations of ED50 and ED10 and possibly LOAEL and NOAEL in life-cycle impact rankings. The experts recognized that when these different measures are combined there is a need for correction factors to make them consistent, to correct for data from different species, and to correct for differences in duration of the experimental exposures. The correction factors should be used to steer all values to best estimates as opposed to the most health protective values. It is important to

avoid introducing bias even with methods that are consistent. There is a need to include a larger number of chemicals in studies of correlation among the different measures of toxicity that were discussed at the workshop.

c) *Complexity of multi-chemical comparison and low dose extrapolation:* The complexities of the analysis make it somewhat difficult to understand how the various dose-response methods might compare with respect to the value of information they provide to an LCIA. This problem derives in large part from differences among the methods of low dose extrapolation. LCIA involves comparisons among the life cycles of products not comparisons among the life cycle of chemicals. But even though this leads to comparisons among chemicals, LCIA is more complicated than single chemical comparisons. Moreover, the problem is one of multi-dimensional optimization. This means that one input depends on others. In LCIA we need a way to compare processes and activities that are not always comparable. The need to confront and inform tradeoffs requires the ability to express relative preferences.

d) *Normalization:* Normalization is a critical issue. It is important to determine what to use as normalizing factor. There are differences between normalizing to manmade changes and to natural background. Global warming provides a cautionary example. The discussion of the relative importance of environmental exposure as a cause of human health impacts and hence the relevance to address human toxic impacts in life cycle impact assessment lead to identification of a need for normalization against other types of impact.

Measures of Severity

With regard to measures of severity, the workshop participants made the following specific recommendations.

a) *Relation between severity and potency:* The expert panel noted the value of taking severity into account but hesitated to recommend specific methods to characterize severity. They took this position because the way that potency is addressed will bear on severity but is not yet fully decided in the LCIA literature. In other words, it is necessary and acceptable to address severity, but we need to work out the specifics of potency issues to decide how.

The panel recognized that not using some measure of severity is the same as treating all outcomes as having equal severity. When confronting severity, there may be some simple methods that are more applicable in some cases, such as distinguishing between carcinogens and non-carcinogens, but, such approaches will also mask real differences in some or many cases. On the other hand, the panel expressed concern that explicitly treating severity may imply that we know more than we do and that we can estimate severity better than is actually feasible. We want to avoid this situation.

b) Multiple effects: It may be appropriate to look at more than one outcome. Similar to other approaches, the RfD approach looks for the critical effect, which is the one that may occur at the lowest dose. The panel noted that this involves toxicity testing that is carried out to find the lowest level that produces an effect rather than completing a whole battery of tests at higher doses. This is appropriate when looking for a “safe” dose but may not be appropriate when trying to characterize representative impacts or the DALY burden on a population. More severe or more common outcomes that are not the ones that occur at the lowest dose may also be important. With only critical effects testing it is not possible to

obtain information on more severe outcomes that occur at higher doses. It may be very important to include a very common but not very severe outcome or a rare but very severe outcome. Thus, there is a need to develop methods to address these concerns.

c) Aggregation: There are aspects of toxic impacts that should be aggregated and other aspects that should be kept disaggregated. There could be elements of severity that fit both. In contrast to policy assessors, modeling/methods experts may have different needs in regard to the issue of aggregation.

If we are going to put a valuation on an endpoint, it is important to be transparent. The experts expressed concerns about DALYs because they represent one judgment about the significance of various health outcomes but these judgments may not be those of the target audience and may not have an established empirical basis. Thus, along with DALYs the analyst should report separately intermediate results such as Years of Life Lost (YOL) or years of life disabled. Additional weighting coefficients should be reported in a transparent way.

Acknowledgements

We are grateful to the workshop participants for the time and effort they put into both the workshop and the preparation of this report. In preparing this report T. McKone was supported in part by the US Environmental Protection Agency National Exposure Research Laboratory through Interagency Agreement #DW-89-93058201-1 with Lawrence Berkeley National Laboratory through the US Department of Energy under Contract Grant No. DE-AC02-05CH11231.

References

P Crettaz, D Pennington, L Rhomberg, K Brand, O Jolliet (2002) "Assessing Human Health Response in Life Cycle Assessment Using ED10s and DALYs: Part 1—Cancer Effects," *Risk Analysis* 22(5): 939-946.

D Pennington, P Crettaz, A Tauxe, L Rhomberg, K Brand, O Jolliet (2002) "Assessing Human Health Response in Life Cycle Assessment Using ED10s and DALYs: Part 2—Noncancer Effects," *Risk Analysis* 22(5): 947-963.

W Krewitt, D Pennington, S Olsen, P Crettaz, O. Jolliet (2002) "Indicators for human toxicity in life cycle impact assessment." In Udo de Haes et al (eds) *Toward Best Available Practice for Life Cycle Impact Assessment* (SETAC Europe Press, Brussels, Belgium).

Gaylor D.W. and L. Swirsky Gold (1998) "Regulatory cancer risk assessment based on a quick estimate of a benchmark dose derived from the maximum tolerated dose," *Regulatory Toxicology Pharmacology* **28**(3), 222-225.

L. Swirsky Gold, D. W. Gaylor and T.H. Slone (2003) Comparison of cancer risk estimates based on a variety of risk assessment methodologies," *Regulatory Toxicology and Pharmacology* **37** 45–53.

L. Swirsky Gold, T. H. Slone, N. B. Manley, G. B. Garfinkel, B. N. Ames (2006) The Carcinogenic Potency Database (CPDB), http://potency.berkeley.edu/

Supplemental Information--Copies of Meeting Presentations

- Presentation by Olivier Jolliet

- Presentation by Lorenz Rhomberg